Kalkulacja rezerw w ubezpieczeniach majątkowych

Wojciech Otto
Wydział Nauk Ekonomicznych
Uniwersytetu Warszawskiego

55 Szkoła Matematyki Poglądowej
Styczeń 2017

Liczba lat opóźnienia *j* / Rok zajścia szkody *t*	0	1	2	...	*T*-2	*T*-1	*T*
0	$x_{0,0}$	$x_{0,1}$	$x_{0,2}$	$\cdots$	$x_{1,1}$	$x_{1,1}$	$x_{0,T}$
1	$x_{1,0}$	$x_{1,1}$	$x_{1,1}$	$\cdots$	$x_{1,1}$	$x_{1,T-1}$	$x_{1,T}$
2	$x_{2,0}$	$x_{1,1}$	$x_{1,1}$	$\cdots$	$x_{2,T-2}$	$x_{2,T-1}$	$x_{T-2,T}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
T-2	$x_{1,1}$	$x_{T-2,1}$	$x_{T-2,2}$	$\cdots$	$x_{T-2,T-2}$	$x_{T-2,T-1}$	$x_{T-2,T}$
T-1	$x_{1,1}$	$x_{T-1,1}$	$x_{T-1,2}$	$\cdots$	$x_{T-1,T-2}$	$x_{T-1,T-1}$	$x_{T-1,T}$
T	$x_{T,0}$	$x_{T,1}$	$x_{T,2}$	$\cdots$	$x_{T,T-2}$	$x_{T,T-1}$	$x_{T,T}$

- $x_{t,j}$ – liczba (lub wartość) szkód zaszłych w roku *t* i zlikwidowanych w roku $(t+j)$.

Na koniec roku *T* znamy wartości zmiennych losowych $x_{t,j}$ którym odpowiadają w tabeli pola jasne: $D = \{(t,j): t \geq 0,\ \ j \geq 0,\ \ (t+j) \leq T\}$,

Nie znamy natomiast wartości zmiennych losowych $x_{t,j}$, którym odpowiadają w tabeli pola zacienione: $P = \{(t,j): t \leq T,\ \ j \leq T,\ \ (t+j) > T\}$.

- Naszym zadaniem jest prognoza wartości sumy: $R = \sum_{(t,j)\in P} x_{t,j}$

  Prognoza jest ważna – jeśli $x_{t,j}$ to wartości szkód, to owa suma to jest właśnie rezerwa na szkody zaszłe, ale na koniec roku *t* jeszcze nie zlikwidowane.
- Jeśli $x_{t,j}$ to liczby szkód, wtedy prognoza sumy R także daje się użyć do kalkulacji rezerwy. Np. jeśli wartość pojedynczej szkody ma ten sam rozkład bez względu na czas zajścia t i czas likwidacji $(t + j)$, wtedy prognozę wartości szkód uzyskamy mnożąc prognozę liczby szkód przez wartość oczekiwaną pojedynczej szkody.
- Przewidzenie liczby (lub wartości) szkód zaszłych, ale jeszcze nie zlikwidowanych, opierać będziemy na próbie wykrycia prawidłowości kryjących się za danymi w lewej-górnej partii tabeli, i ekstrapolacji tych prawidłowości na prawą-dolną część tabeli
- Najpierw zajmiemy się przypadkiem, kiedy $x_{t,j}$ to liczby, a nie wartości szkód.

Dość sensowne założenie to takie, że zmienne losowe $x_{t,j}$ są niezależne oraz pochodzą z rozkładów Poissona o parametrach $x_{t,j} \sim Poisson(\alpha_t r_j)$, gdzie:

- α_t to oczekiwana liczba wszystkich szkód zaszłych w roku *t*,
- r_j to prawdopodobieństwo iż szkoda (zaszła w dowolnym roku) zostanie zlikwidowana *j* lat później.

Przy takich założeniach sensowne rozwiązanie to estymacja parametrów $\alpha_0, \alpha_1, \dots, \alpha_T$ oraz $r_0, r_1, \dots, r_T$ metodą największej wiarygodności, a więc poprzez maksymalizację funkcji:

$$L(\alpha_0, \alpha_1, \dots, \alpha_T, r_0, r_1, \dots, r_T) = \prod_{(t,j)\in D} \frac{(\alpha_t r_j)^{x_{t,j}}}{\left(x_{t,j}\right)!} \exp(-\alpha_t r_j)$$

Przy ograniczeniu $\sum_{j=0}^{T} r_j = 1$, oraz nieujemności wszystkich parametrów.

Okazuje się, że rozwiązanie daje się sprowadzić do bardzo prostej postaci:

$$\hat{\alpha}_0 = \sum_{j=0}^{T} x_{0,j}, \qquad \hat{r}_T = \frac{x_{0,T}}{\hat{\alpha}_0},$$

$$\hat{\alpha}_1 = \frac{1}{1-\hat{r}_T}\sum_{j=0}^{T-1} x_{1,j}, \qquad \hat{r}_{T-1} = \frac{x_{0,T-1}+x_{1,T-1}}{\hat{\alpha}_0+\hat{\alpha}_1},$$

$$\hat{\alpha}_2 = \frac{\sum_{j=0}^{T-2} x_{2,j}}{1-\sum_{j=0}^{1}\hat{r}_{T-j}}, \qquad \hat{r}_{T-2} = \frac{\sum_{t=0}^{2} x_{t,T-2}}{\sum_{t=0}^{2}\hat{\alpha}_t},$$

$$\hat{\alpha}_3 = \frac{\sum_{j=0}^{T-3} x_{3,j}}{1-\sum_{j=0}^{2}\hat{r}_{T-j}}, \qquad \hat{r}_{T-3} = \frac{\sum_{t=0}^{3} x_{t,T-3}}{\sum_{t=0}^{3}\hat{\alpha}_t}, \ldots$$

...

$$\hat{\alpha}_T = \frac{x_{T,0}}{1-\sum_{j=0}^{T-1}\hat{r}_{T-j}}, \qquad \hat{r}_0 = \frac{\sum_{t=0}^{T} x_{t,0}}{\sum_{t=0}^{T}\hat{\alpha}_t}.$$

Na tej podstawie możemy nasze oszacowanie rezerwy wyznaczyć jako:

$$\hat{R} \coloneqq \sum_{(t,j)\in P} \hat{\alpha}_t \hat{r}_j$$

Przedstawione powyżej rozwiązanie może być z powodzeniem zastosowane także w szeregu innych przypadków. W szczególności tak będzie, jeśli zmienne losowe $x_{t,j}$ są niezależne oraz pochodzą z rozkładów złożonych Poissona o parametrach $x_{t,j} \sim złożony\ Poisson(\alpha_t r_j, F)$, gdzie F to dystrybuanta rozkładu wartości pojedynczej szkody. *Należy tu zastrzec, że tożsamość rozwiązania dotyczy tylko samej punktowej prognozy, natomiast rozkład błędu prognozy będzie już inny.*

Zadanie robi się znacznie ciekawsze, jeśli przyjmiemy że wartości zmiennych $x_{t,j}$ to nie są jedyne informacje, jakie posiadamy.

W szczególności zazwyczaj ubezpieczyciel wie, jaka była w każdym z analizowanych lat liczba jednostek ryzyka (*w ubezpieczeniach komunikacyjnych przyjmuje się że jednostka ryzyka to jeden pojazd ubezpieczony przez rok*).

Przyjmijmy oznaczenia:

- n_t dla liczby jednostek ryzyka ubezpieczonych w roku *t*, zaś:
- λ_t dla oczekiwanej liczby szkód na jednostkę ryzyka w roku *t*.

Mamy więc w rezultacie zależność: $\alpha_t = n_t\lambda_t$, gdzie n_t jest znane, zaś α_t potrafimy oszacować. **Wobec tego możemy pośrednio oszacować średnią liczb szkód na jednostkę ryzyka $\boldsymbol{\lambda_t}$.**

Ważkie wnioski możemy uzyskać obserwując, jak bardzo z roku na rok waha się ten parametr. Oczywiście pewną skalę wahań estymatora $\hat{\lambda}_t$ zaobserwujemy zawsze, nawet jeśli prawdziwy parametr $\lambda_t = \lambda$ jest stały w czasie.

Zachodzi pytanie, z jakich właściwie założeń na temat zmienności parametru λ_t wynika nasza metoda estymacji parametrów α_t oraz r_j?

- Odpowiedź brzmi: z żadnych!

Jeśli bowiem założymy, że parametr λ_t jest stały w czasie, wtedy:

- prawidłowa specyfikacja modelu to założenie, że $x_{t,j}$ są niezależne oraz pochodzą z rozkładów Poissona o parametrach $x_{t,j} \sim Poisson(\lambda n_t r_j)$, co prowadzi do estymacji jednego parametru λ i $(T+1)$ parametrów r_j z jednym ograniczeniem.

Alternatywnie możemy założyć, że:

- λ_t to niezależne zmienne losowe o tym samym rozkładzie z wartością oczekiwaną równą Λ oraz wariancją σ_Λ^2, zaś obserwacje $x_{t,j}$ mają warunkowy (przy danym λ_t) rozkład $Poisson(\lambda_t n_t r_j)$, i są (warunkowo, przy danych $\lambda_0, \lambda_1, \dots, \lambda_T$) niezależne. Prowadzi to do metod kalkulacji rezerw bazujących na teorii zaufania (*credibility*).

Oczywiście można wyobrazić sobie cały szereg innych specyfikacji modelu, prowadzących do odmiennych oszacowań rezerwy na szkody zaistniałe, ale nie-zlikwidowane.

Można także metodami statystycznymi szacować jedynie rezerwę na szkody zaistniałe, ale jeszcze nie zgłoszone, zaś w odniesieniu do rezerwy na szkody już zgłoszone oprzeć kalkulację rezerw na doświadczeniu pracowników działu likwidacji.

Można też te same metody stosować do przypadku, kiedy agregacja czasowa naszych danych to nie lata, ale kwartały czy wręcz miesiące.

Problemów, których nawet nie dotknęliśmy, jest mnóstwo.

Jest co studiować!

Dziękuję za uwagę. Kontakt: wotto@wne.uw.edu.pl